ppi radically pragmatic

The Economic Impact of Data: Why Data Is Not Like Oil

Dr. Michael Mandel
Chief Economic Strategist
Progressive Policy Institute

For Presentation at
Data + Privacy Asia Pacific
Sydney, Australia
July 12, 2017

EXECUTIVE SUMMARY

The saying "data is the new oil" is at times referenced by analysts working to assess whether our increasingly digital and data-driven world generates positive impact for our economy and society. However, this saying is imprecise. Data should not be compared to oil – it is not a scarce commodity, is nonrival, and cannot be monopolized.

With regards to privacy, the analogy further weakens. While regulations for traditional commodities like oil seek to protect individual rights to ownership of resources (an individual's oil), the same regulations for the data-driven sector can have negative impact on the economy overall. This is because, when it comes to data, economic value creation is driven by the analysis of data in conjunction with other information. Thus, laws that quite rightfully protect individual rights to data can be at odds with innovation and economic growth.

This tension will get stronger in the coming years as data becomes more important to the economy. Therefore, in this paper, we explore the economic impact of data in the context of privacy regulations, in order to understand how we can make privacy and data-driven industry work together to fuel economic growth. We find that:

• Data has value across borders and across industries. The value, however, is different from the value attributed to physical commodities such as oil.

• The analysis conventionally used to assess the value of physical commodities does not effectively capture the value of data. Unlike physical commodities, data can be reused, is not scarce, cannot be controlled and monopolized by a small number of owners, and has little inherent value alone (without being analyzed). These characteristics affect the design of privacy rules.

• Data-intensive industries have much faster productivity growth than physical industries. In the United States, productivity growth in the digital sector averaged 2.7 percent between 2000 and 2015, compared to 0.8 percent in the physical sector. Moreover, the data-intensive digital sector is creating more jobs. Since the peak of the last business cycle in December 2007, hours worked in the digital category rose 9.6 percent, compared with 5.6 percent on the physical side.

• As physical industries become more data intensive, new jobs are created for middle-skilled workers. For example, in retailing, the shift to data-rich ecommerce since 2007 has created 397,000 jobs in electronic shopping and fulfillment centers in the United States, while only costing 76,000 jobs in brick-and-mortar retailing. Ecommerce giant Alibaba has a delivery and warehouse network in China that employs more than 1.7 million people. Ecommerce in India could generate 12 million net new jobs over the next decade.

• We show how data affects individual interactions with education and training, job matching, consumer purchases, social capital, and exports, among other areas.

• Finally, we identify initial implications of the new economics of data. In particular, we suggest four principles for "growth-friendly" privacy regulation.

The Economic Impact of Data: Why Data Is Not Like Oil

JULY 2017

Dr. Michael Mandel
Chief Economic Strategist
Progressive Policy Institute

INTRODUCTION

Privacy is an important human value. Legislation such as the Australian Privacy Act, the Hong Kong Personal Data (Privacy) Ordinance, and the Singapore Personal Data Protection Act give individuals more control over how their personal data is handled. Such legislation is essential for ensuring that people will be treated fairly and in a nondiscriminatory fashion in today's data-rich world.

However, laws that quite rightfully protect individual rights to data can be at odds with innovation and economic growth. This tension will get stronger in the years ahead as data becomes more important for growth.

In this context, privacy policy will increasingly be shaped by assumptions about the economics of data. In particular, it is common for advocates of strong privacy rules to liken data to oil, a valuable natural resource that exists in limited supply; is essential for most industrial sectors; and that, once consumed, cannot be replaced[1]. This simple economic analogy suggests that rigid privacy rules are needed to make sure individuals get to reap the economic value of their data, just like landowners get to benefit from the oil found under their land.

However, the analogy between data and oil is not only wrong but completely distorts the privacy discussion. Oil is the quintessential physical commodity. The world's oil reserves were laid down millions of years ago, and no more are being made today. Oil companies spend a hundred billion dollars per year to

lay claim to limited fossil fuel reserves which, even today, are essential for economic growth. And the value of a barrel of oil depends on its scarcity.

By contrast, data is not a scarce commodity that needs to be protected by privacy rules. Instead, data is being generated at an exponentially rising rate that is far beyond the capability of any company or group of companies to collect or control [2]. Moreover, unlike a barrel of oil, a piece of data has no value by itself. Instead, creating economic value requires analyzing data in conjunction with other information.

And perhaps the most important way the data-oil analogy distorts the privacy debate: Unlike oil, data is nonrival. That means data can be used multiple times without inherently diminishing its value. Imagine the benefits to the economy if more than one person could consume the same gallon of gasoline!

Indeed, the willingness of individuals and businesses to duplicate and share data is at the heart of the data-driven economy. For example, individual patient data about medicines and medical outcomes can be reused to help determine which medicines work better for what conditions, to help future patients. The same online course can be viewed by multiple students, who can all potentially extract the same benefits. Data about students and performance today can be used to improve education in the future. Data about job postings can be reused by multiple job-seekers to determine where the best jobs are. To the extent that privacy policies discourage this sharing of data, they may have unforeseen economic implications.

> Unlike oil, data is non-rival. That means data can be used multiple times without inherently diminishing its value.

The purpose of this paper is to go beyond the data-oil analogy and develop a more modern understanding of how the economics of data affects privacy rules. Data certainly has value across borders and across industries. That value, however, is different from the value attributed to physical commodities.

The first section makes the argument that the analogy between data and physical commodities such as oil does not effectively capture the value of data. Unlike physical commodities, data can be reused, is not scarce, cannot be controlled and monopolized by a small number of owners, and has little inherent value alone (without being analyzed). These characteristics affect the design of privacy rules.

The second section describes recent research linking data to economic growth, job creation and the reduction of inequality. We show new evidence that industries using data more intensively have faster job growth and pay higher wages. Moreover, as physical industries become more data intensive, wages rise and new jobs are created.

The third section explores several mechanisms by which data, properly applied, can provide benefits to individuals, businesses, and the overall economy. Data affects individual interactions with education and training, job matching, purchases, and exports, among other areas. As a result, the economic analysis of data needs to be rethought.

The fourth section of the paper identifies initial implications of the new economics of data. In particular, we suggest four principles for "growth-friendly" privacy regulation.

WHY DATA IS NOT OIL [3]

Not surprisingly, there is extensive literature on the economic impact of privacy and data protection rules. For example, in 2013, the European Center for International Political Economy (ECIPE) assessed the cost of the then-proposed European General Data Privacy Regulation. Their study found that:

> If the "right to be forgotten" rule is added, the regulation could cause a GDP decrease for the EU of -1.5 percent to -3.9 percent, and welfare loss of 4,566 USD (3,512 euro) per household.4

This excellent study examined the proposed data privacy regulation as a disruption in trade that affected the economy, just like a disruption in trade in oil, cars, corn or some other physical commodity would disrupt the economy. In effect, the authors are studying data by using the existing economic tools for analyzing the conventional economy of goods and services. Similarly, other studies have looked at the cost of data regulation and forced localization, once again in the context of the impact on a conventional economy of goods and services.[5]

The conventional approach to economic analysis treats individuals and businesses as owning scarce commodities such as labor, homes, machines, natural resources such as oil, and so forth. These economic actors trade their scarce commodities for money, to produce outputs, which themselves have to belong to someone. By assumption, production of goods and services has to balance with the various uses of those goods and services in the public and private sectors—not just on the national level, but on the regional and global levels as well.

The balance sheet approach to the economy has served economists well for a hundred years. Why? Because it's well suited to physical commodities such as oil and corn. If one grows corn, the supply of corn in bushels is equal to the bushels of corn eaten by people, plus the bushels of corn consumed by animals, plus the bushels of corn used by industry. Corn supply and demand balance. The same is true for automobiles, steel, health care, and potato chips—the supply and demand balance each other.

The numbers all add up.

In a balance sheet economy, restrictions on data make a difference, but not an enormous one. In a balance sheet economy, privacy rules for personal information may affect advertising, but they don't seem to affect the supply and demand of automobiles, corn, and steel very much.

But here's the rub. As data becomes more important to the global economy, it increasingly bends and perhaps breaks the assumptions underlying the balance sheet view of the economy.

For one, many important services in the data-driven economy are not matched by a monetary transaction. Consumers don't pay to use Facebook or Wikipedia, or to play many mobile games.

TABLE 1: The Policy-Relevant Differences Between Data and Oil

OIL	DATA
Fixed supply, created millions of years ago. We can improve our discovery techniques, but the supply is still limited.	Exponentially increasing supply, with new types of data being created every day.
Because its supply is fixed, oil can be easily controlled by a small number of players, allowing them to drive up the price.	The supply of data is soaring in both volume and type so rapidly that it cannot be controlled or monopolized. Value is generated by new types of analysis rather than ownership of the data.
Unused oil in the ground has a value, set by supply and demand.	Unused data, by itself, has uncertain economic value. Its value depends on how it is combined and used with other data.
Once a barrel of oil is refined and consumed, it is gone.	Data can be duplicated, shared, and reused.
Oil is the single biggest commodity traded on international markets, with the nationality and location of oil reserves and extracted oil tracked very closely. Every barrel of oil that is exported is one less barrel to be consumed at home.	Data can be "exported" to another country without reducing the amount to be used at home. For that reason, it is better to speak of global connections rather than exports and imports of data.

But increased importance of data poses an even more fundamental challenge to the balance sheet view of the economy.

Remember that data is nonrival. That means, unlike oil or corn or cars, data can be duplicated and shared at a relatively low cost, so the production of data does not have to be balanced with the uses of that data.

On one level, we've always known this. The data that describes how to build a computer or an automobile can be used over and over again. And it can be transferred from one country to another without losing any of its value. That's why diffusion of innovation is one of the most potent economic forces.

> Data can be duplicated and shared at a relatively low cost, so the production of data does not have to be balanced with the uses of that data.

For example, a video of a cooking class can be shot and edited once, and then duplicated and shared as many times as there are potential students interested in learning that technique. Similarly, you can go online and find information about how to build a house, summarized, duplicated and shared costlessly. Furthermore, if one application passively captures GPS data from an individual's mobile device, for instance, a mapping application, it does not preclude another application, such as a restaurant search application, from capturing that very same data.

Let us take this a step further. Because data can be duplicated nearly costlessly, there can never be a scarcity of data in the same way there can be a scarcity of food or oil. Quite the contrary: There is an exponentially increasing amount of data available to everyone.

Indeed, unlike oil, the underlying raw data actually has very little value, but instead the thing that is creating the value – the thing that creates economic value where there was none – is the collection and analysis of data and the ability to build useful products and services on top of it.

This is not simply a pedantic sidelight—it goes to the very heart of how we think about production, consumption, and trade. Consider the fundamental identity of international trade.

> Domestic purchases = Domestic production + imports - exports

For steel, corn, and automobiles, this identity is so obvious it goes without saying. The amount of steel available to be purchased in Australia, say, is equal to the amount of steel produced plus the imports of steel minus the exports of steel.

But this identity breaks down completely in the data arena. Suppose a mobile app is developed in Australia or Singapore or Korea or Japan, and then exported to the United States. Unlike steel, that same app is still available in the home country, just as if it had never been exported. When it comes to data, exports don't take away from the amount of data available at home.

If we are to understand the impact of data and privacy regulations—and especially if we are to understand the fundamental connections among Asia-Pacific countries such as Australia, Singapore, Korea, Japan, the Philippines, the United States, Canada, and Mexico—the balance sheet concepts of production, consumption, and trade need to be rethought. We measure output of goods and services, but we don't measure the data created and how it is put to use. We measure how much consumers spend, but not how the value of their time changes. We measure the flows of goods and services among the different countries, but we don't measure the duplication and sharing of intangibles and how effectively they are used.

MORE DATA LEADS TO BETTER PRODUCTIVITY AND WAGE GROWTH

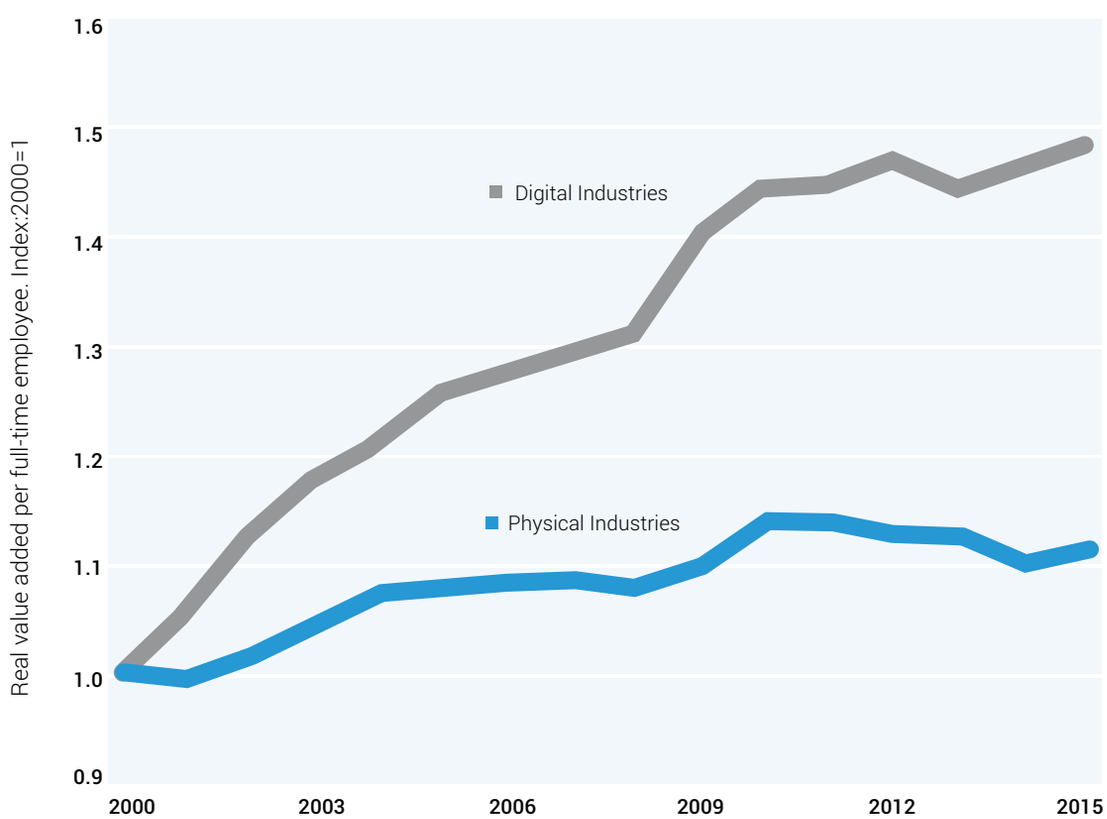
In this section we describe new research showing how more data-intensive industries have faster productivity growth, stronger job creation, higher wages, and reduced income and geographic inequality. Many developed countries have been suffering from slow or sluggish growth in recent years. But it's increasingly clear that better use of data is the key to faster growth.[6]

On a macroeconomic level, we note first that the private sector can be divided into two types of industries: digital industries and physical industries.[7] The digital sector includes industries such as tech, telecom, entertainment, financial services, and professional services. The physical sector includes industries such as manufacturing, construction, transportation, healthcare, education, and retail.

In the United States, the digital sector, which represents 30 percent of private sector output and 25 percent of private sector employment, makes 70 percent of US private sector information technology investment. The physical sector, which represents 70 percent of private sector output and 75 percent of private sector employment, makes 30 percent of US private sector IT investment.

Clearly the digital sector is much more data-intensive. Moreover, it has much higher productivity growth. According to our calculations, the digital sector has averaged 2.7 percent productivity growth since 2000, while the physical sector has averaged only 0.8 percent annual productivity growth (Fig. 1).

FIGURE 1: Productivity Growth: Digital vs. Physical



Sources: Bureau of Economic Analysis, author calculations

What about job creation? Surprisingly, the digital industries have had stronger job creation in recent years than the physical industries. Since the peak of the last business cycle in December 2007, hours worked in the digital category rose 9.6 percent, compared with 5.6 percent on the physical side. If health care is excluded, hours worked in physical jobs rose only 3 percent.

As physical industries digitize and use more data, they are creating more jobs and raising wages. For example, in retailing, the shift to data-rich ecommerce since 2007 has created 397,000 jobs in electronic shopping and fulfillment centers in the United States, while only costing 76,000 jobs in brick-and-mortar retailing. Moreover, preliminary results show that the pay in the ecommerce jobs is on average 30-40 percent higher than in brick-and-mortar retail.[8]

And it's not just the United States. Ecommerce giant Alibaba has a delivery and warehouse network in China that employs more than 1.7 million people.[9] And a report from HSBC suggests that ecommerce in India could generate 12 million net new jobs over the next decade, over and above the lost jobs in brick-and-mortar retail.[10]

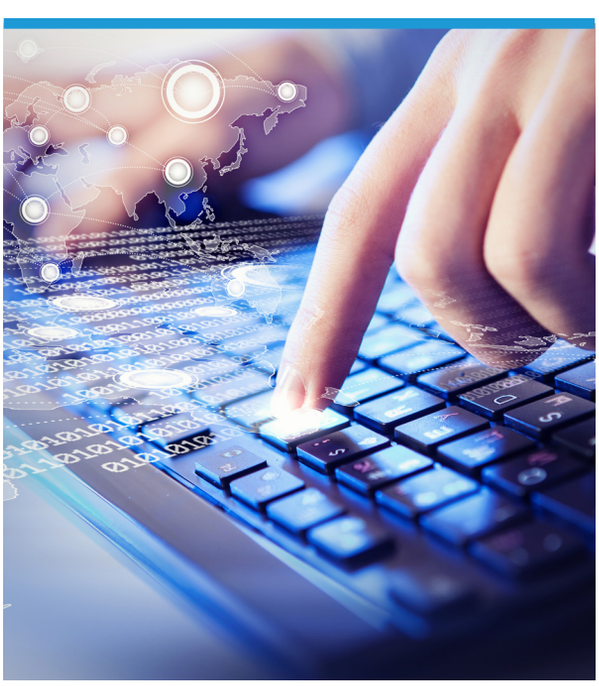
The key is that data allows improvement of the distribution system, which in turn reduces the number of hours each week that consumers have to shop. This increase in productivity creates new well-paying jobs in fulfillment centers, while making consumers better off.

At the heart of this gain is the use of data to create new products and services that benefit both consumers and workers.

Another issue that has troubled many countries is rising inequality. That comes in two forms: income inequality and geographic inequality. Income inequality is the gap between high and low earners, while geographic inequality is the tendency for economic growth to concentrate in relatively few regions. Preliminary research suggests that income inequality will narrow as more physical industries become digitized and wages rise in those industries.

> At the heart of this gain is the use of data to create new products and services that benefit both consumers and workers.

Moreover, jobs created by digitization of physical industries seem to be geographically better distributed. The top ten states for ecommerce job growth in the United States include Pennsylvania, Indiana, Georgia, Tennessee, and Ohio. These states, which were hit hard by the decline of manufacturing, are now being helped by the rise of ecommerce. Similarly, ecommerce and other uses of data to upgrade and digitize physical industries can create jobs in Asia-Pacific countries outside of the major cities.



A TYPOLOGY OF THE ECONOMIC BENEFITS OF DATA

As noted earlier, discussions about privacy rules often are based on the assumption that data is similar to a physical commodity like oil. In some sense, this approach is only natural: We have extensive experience with physical commodities, but widespread use of data is something new. In particular, the reusability or "non-rival" quality of data creates many new opportunities for generating economic growth.

Moreover, raw data by itself has little value, whether in small or large quantities, without an innovative idea about how to analyze and use it. Governments and nonprofits, for example, often have access to large amounts of raw data about individuals, but don't know how to use it. It is actually the application of data that creates the economic value.

In this section we briefly explore several mechanisms by which data, properly applied, can provide benefits to individuals, businesses, and the overall economy. These mechanisms include the use of data to improve education and training for young people and adult learners; better job matching for workers; more information for consumer purchases; increased investment in social capital; higher productivity and wages; better opportunities for exports, especially for small and medium enterprises (SMEs); improved government performance; and global macroeconomic gains.

TABLE 2: The Economic Benefits of Data: Some Policy-Relevant Mechanisms

MECHANISM	SUMMARY
Education and learning	Personalized online education and training will be essential for helping prepare young workers and retool mid-career workers at an affordable cost. But improving outcomes for online education will require intensive use of personal data.
Job matching	Both developed and developing countries can do better at matching job seekers with good opportunities. Such a system will require extensive use of data on skills, personality, and outcomes.
Consumer purchases	Consumers are benefiting from information not just on the price but on the quality of products and services, based on personal reviews.
Social capital and social networks	The increase in social capital through social networks can boost innovation and economic outcomes.
Productivity and wages	Data-intensive industries have faster productivity growth, higher wages, and stronger job creation. In the United States, for example, ecommerce has created almost 400,000 decent-paid jobs since 2007, while brick-and-mortar has lost less than 80,000.
Exports	Access to data makes it easier for small and medium enterprises to take advantage of foreign markets.
Improved government performance	Increased use of data can improve government performance and civic engagement.
Global macroeconomic gains	Cross-border data flows allow intangible capital such as scientific knowledge and management technique to spill over to other countries, boosting global macroeconomic performance.

Source: Progressive Policy Institute

Education and Training

Educating and training workers is key for growth, both in developing and developed countries. Formal education by itself is a significant portion of GDP almost everywhere, amounting to 5.6 percent of GDP in Australia and 5.9 percent of GDP in Korea, based on OECD statistics.[11] But that's only a portion of the education and training expenditures. As innovation accelerates, older workers will have to learn new skills in order to remain productive and employable.

Unfortunately, training and education are highly inefficient industries today, conducted mostly in person via methodologies—such as large lectures and textbooks—that have been around for decades. Countries such as Singapore and Hong Kong, which have top-ranked educational systems, have faced the question of how to instill entrepreneurial and innovative traits in students.[12] It's impossible to scale up the current education/training model to accommodate the needs of the 21st century.

So far we've had some small steps in this direction, in the form of MOOCs (Massive Open Online Content). As of 2016, 58 million students were enrolled globally in MOOCs.[13] In total, there are almost 7000 MOOCs over 700 universities around the world, including China, Latin America, Europe, the Middle East and North Africa. India has launched the MOOC platform SWAYAM.

However, today's online education is far short of what will be needed in the future, both in scale and scope. MOOCs are one-size-fits-all, rather than being designed to adapt to the individual learners. That's especially important for older workers, who are typically less attuned to online learning.

What's needed is a breakthrough in personalized online education and training. In that way, the power of data can be used to create online courses that then can be used by thousands or millions of learners to reduce costs, while being customized to suit the individual based on their background and needs.

> However, today's online education is far short of what is needed in the future, both in scale and scope

Achieving this personalized education will require analysis of large amounts of data about the link between learner traits and optimal use of educational resources and strategies.[14] Meeting this goal could have an enormous economic impact, since recent research suggests that online modes of education offer the possibility of reducing the cost of education and increasing competition.[15] One 2016 report notes that: "Our results suggest that by increasing competitive pressure on local schools, online education can be an important driver of innovation and productivity in U.S. higher education."[16]

Better Job Matching

Throughout this economic recovery, there continues to be questions about the efficiency of national job markets. Companies complain around the world that they cannot get enough skilled workers. Workers complain that they cannot find jobs that make use of their skills and talents. As of February 2017, the Australian Bureau of Statistics reported 187,000 job vacancies in that country. In

Mexico, Indeed.com, a global job search engine, reports more than 200,000 job postings as of June. In the United States, the Bureau of Labor Statistics reported more than 6 million job openings in April. Hong Kong, with obviously a much smaller economy, reported 79,000 job vacancies as of March 2017.

There is a growing use of online job postings by employers, which benefit workers. One economic study showed that Internet job search (IJS) reduces individual workers' unemployment durations by 25 percent.[17] Such online job postings make it easier for workers to identify potential jobs, not just in their local area but in other parts of their country.

But online job postings are not enough to bring the labor market into the 21st century—the other side of the market is important as well. Employers need to be able to systematically search for workers who will be successful at their companies. We should be accumulating data about what type of skills and experience make workers suited for which types of jobs, and that data should be available to all employers. Moreover, done right, this database could support diversity and mobility by giving companies access to the full range of potential employees who could be successful. Does this database exist yet? Companies such as LinkedIn have scratched the surface. But the potential is enormous.

> One economic study showed that Internet job search (IJS) reduces individual workers' unemployment durations by 25 percent.

Consumer Purchases

Consumers have access to far more data than ever before about their potential purchases. If they have access to a computer, they have the ability to compare prices and models, not just for their local stores but for merchants around the country.[18]

Moreover, consumers can compare not only prices but perceived quality as well. Most major buying sites have a provision for buyer reviews, with some sites, such as TripAdvisor, being heavily devoted to such reviews. Indeed, the mere existence of reviews for a product—showing that it is being purchased and used by others—serves to stimulate purchases.

This data on the perceived quality is extremely important, not just for existing products but for the introduction of new innovations.[19] In the past, new products and services seeped into the market slowly, as consumers waited to hear about them from trusted friends. But today the process of the introduction of new products and services can be accelerated by use of data.

This data on the quality of products and services, while strictly speaking not personal

information, is a richer resource if we know something about the characteristics of the people writing the reviews. In theory, such data can drive a systematic improvement in the quality of products and services across the whole economy.

Social Capital and Social Networks

According to one source, individuals globally spend an average of 2 hours per day on social networks such as Facebook, up half an hour since 2012.[20] This time does not show in GDP, because no money exchanges hands. It has economic value nevertheless.[21]

A variety of studies have attempted to estimate the contribution of social networks to consumer welfare or GDP. For example, one 2012 study estimated that unmeasured data consumption potentially added a half percentage point to GDP growth.[22] Other studies have come to similar conclusions.[23]

> …the most-connected metro regions had more than double the job growth of the least-connected metro areas.

But there's another aspect of social networks that gets far less attention: The creation and accumulation of *social capital*.[24] Social capital is defined as the aggregate of the social and economic resources available to people through connections with other individuals. It is generally accepted that social capital, like physical capital, adds to the productivity of an economy. Moreover, social capital is not owned by a single person or organization, but rather is accessible to every member of the social network.

It should be true that cities, regions, and countries with denser social networks have stronger growth.[25] One study found that

…the most-connected metro regions had more than double the job growth of the least-connected metro areas.

The top quintile of metro areas, ranked by their index of connectedness, had an average job growth of 8.2 percent in the four years from 2010 to 2014. The bottom quintile of metro areas, ranked by their index of connectedness, had an average job growth of only 3.5 percent.

Further research is needed on a wider variety of social networks, but it appears that social networks are a plus for growth.

Exports

It's a well-known fact that global trade is dominated by large companies. In the United States, companies with more than 500 employees account for two-thirds of exports and imports, even though they only account for 3 percent of exporters and importers.[26] Similar statistics are true for other countries as well.

One of the biggest barriers blocking small and medium enterprises (SMEs) from global trade is a lack of knowledge about opportunities. It's a vicious circle: Without personal connections overseas, it's very difficult for SMEs to learn about export and import opportunities. But the personal connections don't get built without the contacts overseas.

But cross-border data flows open up new opportunities for SMEs to trade overseas.[27] It's easier to locate customers, and easier to build supply chains, across national borders.

Government Gains from Data

Many governments are working towards an "open data" policy, where they allow other organizations access to government-collected data. But governments, and more generally society at large, benefit from data generated in the private sector as well. A recent review of the economics of privacy noted that.[28]

In particular, the benefits arising from individuals sharing their information, because of advances in data mining, may be enjoyed by society as a whole. For instance, aggregation of online searches can provide early alerts for epidemics…or unveil unexpected interactions between pharmaceutical drugs.

In a broader sense, data may also be crucial for fostering civic engagement.

Global Macroeconomic Gains

Finally, one of the biggest benefits of cross-border data flows is the spillover of intangible capital from one country to another. It is well known that intangible capital, such as research and development, design, management capabilities, and financial knowledge, comprises a key motive force of the global economy.

These intangible spillovers, channeled through cross-border digital flows, allow one country to benefit from another country's technological and institutional advances. So, when the United States leaps ahead in an area such as mobile apps, these innovations are very quickly transmitted to other countries, where they create jobs.

A big question is whether privacy and data protection rules hinder the cross-border spillovers of intangible capital. Unlike trade in goods and services, we don't measure cross-border data flows very well. So we won't see the direct impact of privacy and data protection rules on these cross-border data flows. This is an area where more research is greatly needed.

IMPLICATIONS AND PRINCIPLES

In the previous sections of this paper, we first described an economic framework in which data is a resource that can be duplicated, shared and utilized across the entire economy. Then we drilled down to a typology of mechanisms by which data can increase economic growth and individual welfare.

Indeed we also note that the ability to share data and to use it for multiple purposes is precisely what powers economic growth. The position that data is like oil is precisely wrong, and not a good model for designing privacy rules. In fact, in the data economy, individual privacy can be at odds with economic growth. Individual privacy decisions can dramatically hinder the feedback loop that data aggregation and analysis creates, thus resulting in subpar innovation in the economy.

To address an often-raised question, the sheer aggregation of existing data by itself does not create antitrust issues.[29] Data is nonrival, so the mere fact that one company collects a particular type of data doesn't stop other companies from doing the same thing. For example, data on energy usage by household and by time of day is very important for designing the optimal smart electronic grid. But any electric provider has sufficient information, if they have the right analytic tool.

> Individual privacy decisions can dramatically hinder the feedback loop that data aggregation and analysis creates, thus resulting in subpar innovation in the economy.

Or consider the education system. Most countries need to find new ways to make their education system more effective, both for the young and for mid-career job switchers who are trying to retrain. The best solution would be personalized online education, which fits the needs of the individual learner.

But developing effective personalized online education would require analysis of a large amount of data linking de-identified personal characteristics to both education and work outcomes. What's more, the biggest improvements are likely to come from outside the current education and training system, since today's big educational institutions have a vested interest in minimal departures from the status quo. Given that data can be duplicated, successful challenges to the current system could spread very quickly, because new firms can easily and cost-effectively capture relevant data in a way that is nonrival with other educational firms. In order to allow for the strongest outcomes for individuals, privacy rules would need to allow and encourage data aggregation, collection and analysis.

From that perspective, we can identify four principles for fostering data-driven growth and innovation:

Focusing opt-in consent obligations on the riskiest cases can help promote innovation while giving individuals choice over their privacy standards:
We should focus opt-in consent obligations on the cases where there is the greatest concrete risk to the individual, where we need them to pay attention to make sure the particular use is something in which they want to engage in. For other uses that don't create this kind of risk, we should consider other models – which could include default rules that allow particular

uses (such as fraud prevention) or that enable people to opt out.

If users must opt into their data being shared for individual apps, the amount of analysis paralysis would be overwhelming: how can a user identify which apps and technologies they should give data to that will best build individual economic benefit? As we have seen in the case of improving the education and training system, some of the most important economic uses of data may be indirect and nonobvious. We don't actually know right now which individual characteristics are important for personalizing education. But we do know that once we figure it out, the potential gains could be enormous.

Additionally, by decreasing the individual data points available across different companies, the value of data in the aggregate decreases, which could decrease the innovation potential for a given sector. This is especially the case when you consider the use of data for secondary purposes – if a user doesn't know that their data can be used to find a cancer cure, how can they know to opt-in? If a company, knows the data is valuable for their research, how much time will they lose chasing consent from individual users?

The de-identified linking of records can jointly encourage strong privacy practices and innovation:

Privacy rules should take into account that a piece of data by itself has no value—its value comes from its analysis in conjunction with other data. That suggests privacy rules should allow, by default, de-identified linking of records across different data domains.

In this case, statistics and data privacy go hand in hand. The more de-identified data that is available for processing, the more privacy preserving, and statistically accurate the dataset is. The danger comes not from too much data, but from too little. Ultimately scale works in favor of privacy.

Cross-border data flows result in economic growth and benefits to individuals:

Privacy rules should acknowledge the economic benefits of cross-border flows of data. As we saw earlier, cross-border data flows are better described as "connections" rather than exports and imports. Unlike oil, when data crosses a border, it is still present and usable in the original country. Moreover, data flows are generally bidirectional, so setting up a connection gives residents of one country access to data from another country. Indeed, some of the biggest benefits of data come from cross-border spillovers. We don't have the right statistical system to track them yet, but that's essential for growth in the global economy.

Privacy and innovation can and should go hand in hand:

Fourth, and most important, privacy rules should explicitly specify that data regulators should give weight to innovation as a secondary but important goal of policy. The model is the United States Federal Reserve, which has a dual mandate of price stability and full employment. Data regulators should be given a dual mandate as well.

TABLE 3: Privacy Principles for Fostering Data-Driven Growth

PRINCIPLE	ECONOMIC BENEFIT
Focusing opt-in consent obligations on the riskiest cases can help promote innovation while giving individuals choice over their privacy standards.	Data gets more valuable as it is shared. The riskiest cases should be opt-in. But, otherwise, opt-out will encourage individuals to be selective about what information is kept private.
The de-identified linking of records can jointly encourage strong privacy practices and innovation.	Some of the biggest economic gains will come from jointly analyzing information from multiple domains, such as education and work outcomes.
Cross-border data flows result in economic growth and benefits to individuals.	Cross-border data flows facilitate the transmission of intangible capital from one country to another, which is one of the most powerful forces for economic growth. Moreover, movements of data across borders still typically leave the original data in place.
Privacy and innovation should go hand in hand.	As new forms of data and data analysis become available, data regulators should acknowledge the importance of promoting growth and innovation.

CONCLUSION

We are in the early years of applying data to the economy, especially in physical industries such as retail, manufacturing, and education. Broader applications of data are likely to be crucial for generating both economic and job growth in the years ahead. Now is the right time to make sure we have the right economic framework to understand how we can make privacy and data-driven industry work together.

References

1. "The world's most valuable resource is no longer oil, but data," *Economist,* May 6, 2017. See also "Why Data is the New Oil," *Fortune*, July 11, 2016, and "From fintech to techfin: data is the new oil," *The Asian Banker,* May 15, 2016.
2. According to Cisco's latest forecast, global data traffic is expected to triple over the next five years. http://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html.
3. The section builds on Michael Mandel, "Moving beyond the balance-sheet economy" in *Policy Choices For A Digital Age*, Friends of Europe, June 2017.
4. "The Economic Importance of Getting Data Protection Right: Protecting Privacy, Transmitting Data, Moving Commerce, " European Centre for International Political Economy (ECIPE), March 2013.
5. Brendan O'Connor, "Quantifying the Cost of Forced Localization," Leviathan Security Group, June 2015.
Matthias Bauer, Martina F. Ferracane, and Erik van der Marel, "Tracing the Economic Impact of Regulations on the Free Flow of Data and Data Localization," Centre for International Governance Innovation and Chatham House, May 2016.
6. https://hbr.org/2016/01/the-most-digital-companies-are-leaving-all-the-rest-behind.
7. Michael Mandel and Bret Swanson, "The Coming Productivity Boom," March 2017.
8. Michael Mandel, "Update on ecommerce and brick-and-mortar retail jobs" Progressive Policy Institute, May 12, 2017.
9. Alibaba Group, "Alibaba Group Announces June Quarter 2016 Results." August 11, 2016. http://www.alibabagroup.com/en/news/press_pdf/p160811.pdf.
10. Pranjul Bhandari, "India: More jobs per click--the employment potential of e-commerce'," HSBC, August 26, 2016. http://www.hsbc.com/news-and-insight/insight-archive/2016/india-more-jobs-per-click.
11. OECD, *Education at a Glance, 2016*.
12. "Going beyond grades: Evolving the Singapore education system," *The Strait Times,* April 17, 2016.
13. "Monetization over Massiveness: A Review of MOOC Stats and Trends in 2016"
https://www.edsurge.com/news/2016-12-29-monetization-over-massiveness-breaking-down-moocs-by-the-numbers-in-2016.
14. Chris Dede and Andrew Ho. "Big Data Analysis in Higher Education: Promises and Pitfalls," Educause Review, August 2016.
15. David J. Deming, Claudia Goldin, Lawrence F. Katz, Noam Yuchtman ."Can Online Learning Bend the Higher Education Cost Curve?" NBER Working Paper No. 20890, January 2015.
16. David J. Deming, Michael Lovenheim, Richard W. Patterson, "The Competitive Effects of Online Education," NBER Working Paper No. 22749, October 2016.
17. Kuhn, Peter J. and Hani Mansour. "Is Internet Job Search Still Ineffective?" Discussion Paper Series. Forschungsinstitut zur Zukunft der Arbeit, No. 5955.
18. Alberto Cavallo. "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers." NBER Working Paper 22142, 2016.
19. Itamar Simonson and Emanuel Rosen. "What Marketers Misunderstand About Online Reviews" Harvard Business Review, January-February 2014.
20. http://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/
21. http://www.digitopoly.org/2017/06/10/the-value-of-free-in-gdp/
22. Michael Mandel, "Beyond Goods and Services: The (Unmeasured) rise of the Data-Driven economy," Progressive Policy Institute, October 2012.
23. Leonard Nakamura, Jon Samuels, and Rachel Soloveichik, "Valuing 'Free' Media in GDP: An Experimental Approach."FRB of Philadelphia Working Paper No. 16-24, August 2016.
Erik Brynjolfsson and Joo Hee Oh, , "The Attention Economy: Measuring the Value of Free Digital Services on the Internet," MIT Working Paper, 2012.
24. https://www.oecd.org/insights/37966934.pdf.
25. Michael Mandel. "Connections as a tool for growth: Evidence from the LinkedIn Economic Graph." South Mountain Economics, 2014.
26. U.S. Census Bureau, "A Profile of U.S. Importing and Exporting Companies, 2014 – 2015," April 4, 2017. https://www.census.gov/foreign-trade/Press-Release/edb/2015/edbrel.pdf.
27. Ed Gerwin, "The Digital Opportunity: Democratizing Trade for the 99 Percent," Progressive Policy Institute, May 2015.
28. Alessandro Acquisti, Curtis Taylor, and Liad Wagman. "The Economics of Privacy." Journal of Economic Literature, 54(2), 2016.
29. Joe Kennedy, "The Myth of Data Monopoly: Why Antitrust Concerns About Data Are Overblown," ITIF, March 2017.

ppi

The Progressive Policy Institute is a catalyst for policy innovation and political reform based in Washington, D.C. Its mission is to create radically pragmatic ideas for moving beyond ideological and partisan deadlock.

Founded in 1989, PPI started as the intellectual home of the New Democrats and earned a reputation as President Bill Clinton's "idea mill." Many of its mold-breaking ideas have been translated into public policy and law and have influenced international efforts to modernize progressive politics.

Today, PPI is developing fresh proposals for stimulating economic innovation and growth; equipping everyone with the skills and assets that social mobility in the knowledge economy requires; modernizing an overly bureaucratic and centralized public sector; and, defending liberal democracy in a dangerous world.

Progressive Policy Institute
1200 New Hampshire Ave NW,
Suite 575
Washington, DC 20036

Tel 202.525.3926
Fax 202.525.3941

info@ppionline.org
progressivepolicy.org